NETW 741
2995

UNIVERSITEIT GENT

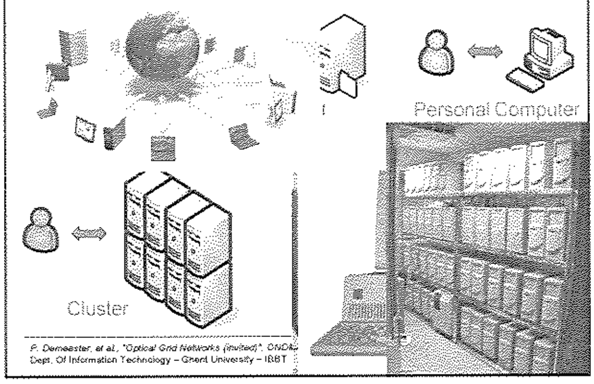
FACULTY OF ENGINEERING

Optical Grid Networks
(invited)

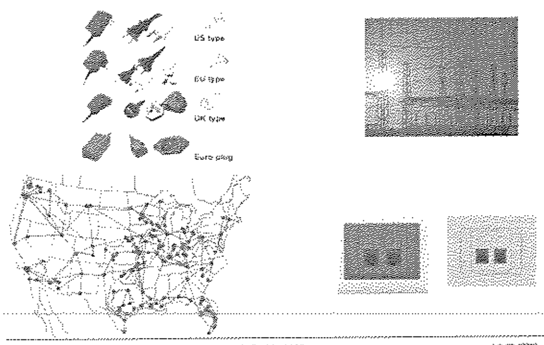
P. Demeester, M. De Leenheer, C. Develder, T. Stevens, J. Vermeir, F. De Turck, B. Dhoedt

11th Int. Conf. on Optical Netw. Design and Modeling (ONDM)
Athens, Greece, May 30th, 2007

Evolution towards grid computing

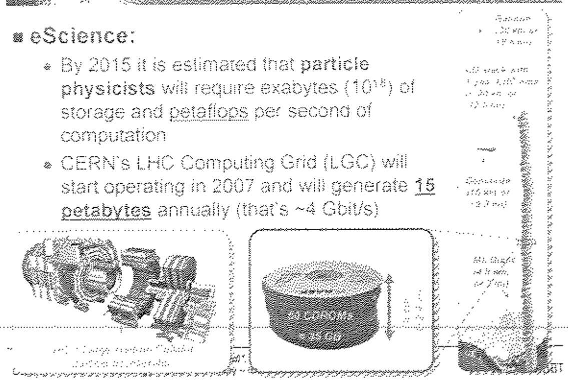


Evolution towards grid computing



Introduction (1)

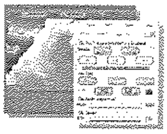
- eScience:
  - By 2015 it is estimated that **particle physicists** will require exabytes (10^{18}) of storage and petaflops per second of computation
  - CERN's LHC Computing Grid (LGC) will start operating in 2007 and will generate **15 petabytes** annually (that's ~4 Gbit/s)



UNIVERSITEIT GENT

Introduction (2)

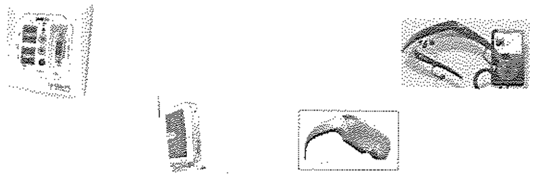
- **Consumer service:**
  - Eg. **video editing**: 2Mpx/frame for HDTV, suppose effect requires 10 flops/px/frame, then evaluating 10 options for 10s clip is 50 Gflops (today's high performance PC: ~10 Gflops/s)



P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT

Introduction (3)

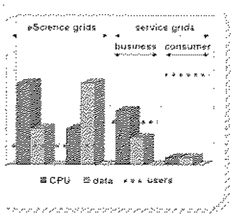
- Move towards (mobile) thin clients
- Limited processing and storage in the terminals
- Use of server based computing/storage



P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT

Introduction (4)

- Grid opportunities ranging from academia over corporate business to home users



P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT

UNIVERSITEIT GENT

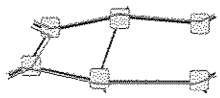
FACULTY OF ENGINEERING

P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT

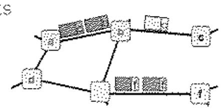
UNIVERSITEIT GENT

Optical Network Architecture

- **Optical Circuit Switching (OCS)**
  - continuous bit-stream
  - pre-established light-paths
  - should be dynamic



- **Optical Burst/Packet Switching (OBS/OPS)**
  - chunks of bits, in bursts/packets
  - forwarding based on header
  - e.g. label switching, GMPLS

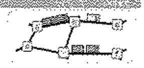


- **Hybrids**

Optical Circuit Switching

- **Pro:**
  - Guaranteed service quality once set-up (cf. reserved lambda), thus fixed latency, no jitter, etc.
  - Fixed signaling overhead, independent of (large) job size
- **Con:**
  - Signaling overhead† not acceptable for relatively small jobs
    - **Requires (complex) grooming if frequent set-up and tear-downs are to be avoided (i.e. if too slow)**
  - Less flexible, dynamic than OBS/OPS, cf. light-path set-up and tear-down

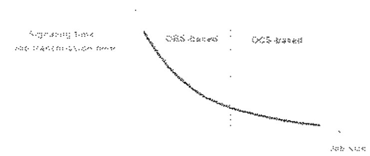
OBS/OPS



- **Pro:**
  - Extremely flexible, dynamic
  - Inherent statistical multiplexing of available bandwidth (over multiple lambdas)
- **Con:**
  - Packet/Burst header processing overhead
    - **Requires job aggregation if job size too small compared to header overhead**
  - Difficult to deliver strict QoS guarantees without 2-way reservation
  - Technology not that mature (hardware)

Hybrid OCS/OBS

- **Choosing between OCS and OBS depends on...**
  - Optical technology (OBS requires faster switches, burst mode Rx/Tx and regenerators, ...)
  - Job sizes:

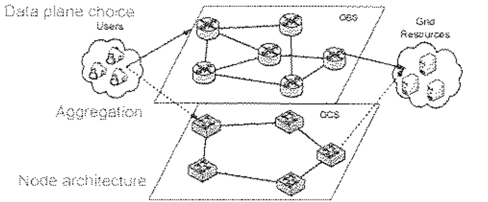


- **Hybrid architectures can offer a compromise**

UNIVERSITEIT GENT

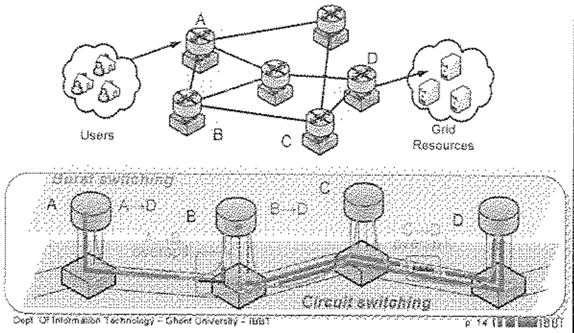
Hybrid OBS/OCS

- **Parallel: choice to either set-up OCS circuit between source & destination, or use OBS**
  - Note: can be overlay, where OBS makes use of OCS connections between OBS nodes



Hybrid OBS/OCS: ORION

- **Overspill Routing In Optical Networks:**



Challenges

- **Grid applications pose challenging demands:**
  - Connections extend to application end-points (rather than traditional network elements)
  - On-demand bandwidth provisioning, both immediate and advance reservations
  - Very dynamic use of end-to-end networking resources
  - Requires (near) real-time feedback for signaling and provisioning
  - Heterogeneous network, computing and storage resources in multiple domains
  - Diversity in holding times and bandwidth granularity
  -

UNIVERSITEIT GENT

FACULTY OF ENGINEERING

Introduction
Network Architecture
Routing
Multiple Domains
Dimensioning
Conclusions

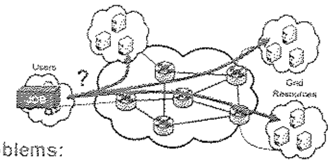
UNIVERSITEIT GENT

Optical Grid specifics

- **Differences with "classical" optical networks or "classical" Grids:**
  - Anycast routing: user generally doesn't care where job is executed
  - Burst starvation: not only network contention, also Grid resource contention
  - Future reservation!: some jobs have very loose response time requirements, others are known long beforehand

Problem Statement

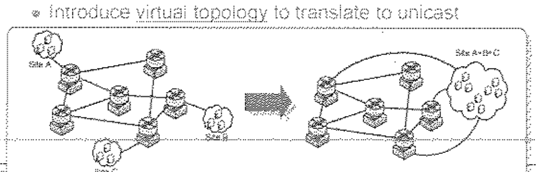
- **Problem:**
  - Given a job submitted by a user to an anycast address
  - Find a set / containing at least one (and preferably one) suitable Grid site location accepting such jobs + a route to reach it



- **Sub-problems:**
  - Routing/deflection strategies
  - Distributed multi-constrained routing algorithms

Anycast SAMCRA

- **Problem:**
  - Incorporation of other metrics than just Grid resource availability leads to a multiple-constraint anycast routing problem
    (unicast multiple-constraint is already NP-complete)
- **Our solution:**
  - Introduce virtual topology to translate to unicast



Anycast SAMCRA

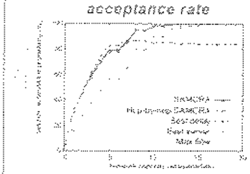
- **Problem:**
  - Incorporation of other metrics than just Grid resource availability leads to a multiple-constraint anycast routing problem
    (unicast multiple-constraint is already NP-complete)
- **Our solution:**
  - Introduce virtual topology to translate to unicast
  - Use a Self-Adaptive Multiple Constraint Routing Algorithm (SAMCRA)
  - Use a novel path ordering avoiding sub-optimality and loops

UNIVERSITEIT GENT

Anycast SAMCRA: results

- Comparison with a shows that even distributed SAMCRA comes very close to (pseudo-)optimal acceptance rate
- Simpler heuristics, taking only 1 measure into account, do not come as close

acceptance rate



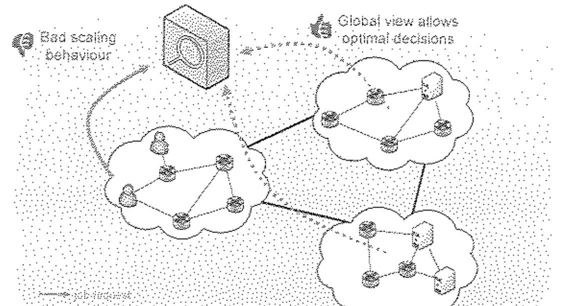
T. Stevens et al., "Distributed Job Scheduling based on Multiple Constraints Anycast Routing", Broadnets 2006

UNIVERSITEIT GENT

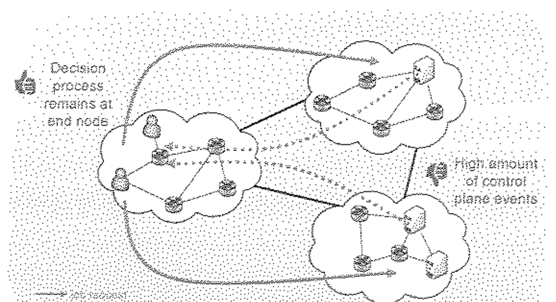
FACULTY OF ENGINEERING

Introduction
Network Architecture
Routing
Multiple Domains
Dimensioning
Conclusions

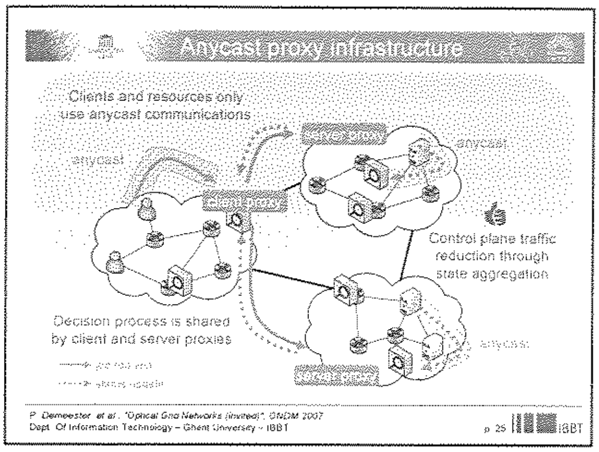
Centralized

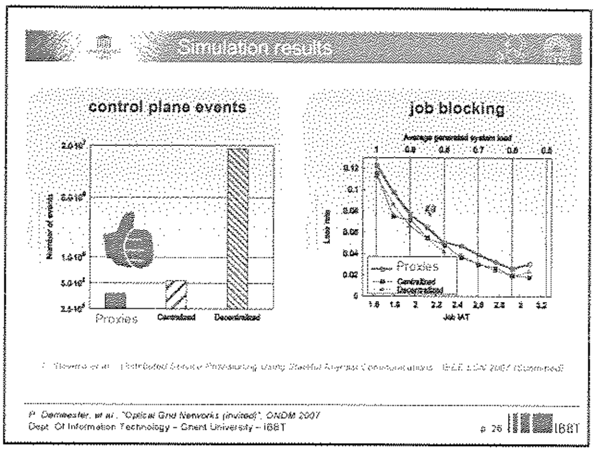


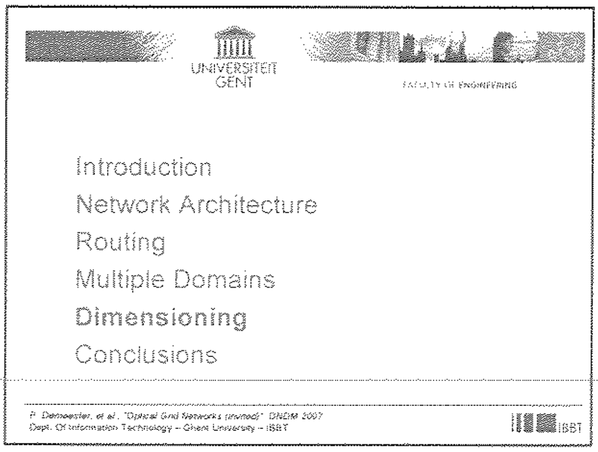
Distributed

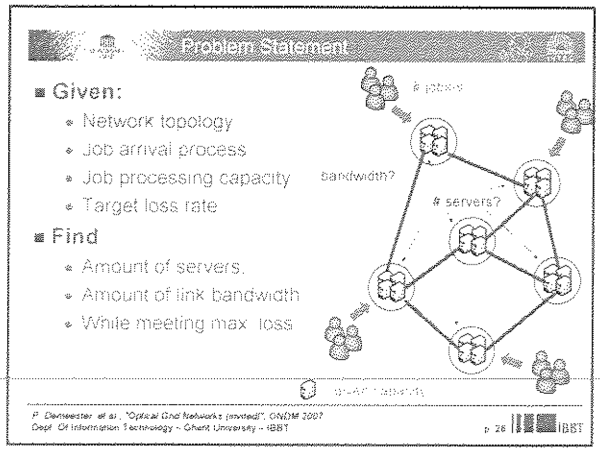


UNIVERSITEIT GENT





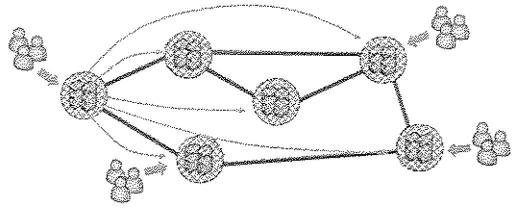




UNIVERSITEIT GENT

Solution

- Phased approach
  - Determine total server capacity (analytical, ErlangB)
  - Determine inter-site bandwidths (simulation)
  - Dimension link bandwidths (=number of wavelengths)

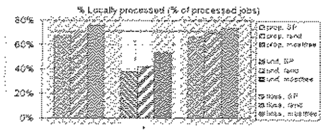


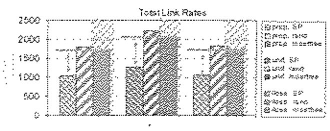
Algorithms

- Step 1: total # servers n follows from ErlangB, distribution among sites (N):
  - *unif*: uniformly distributed among all sites: $n_i = n/N$
  - *prop*: proportional to local arrival rate: $n_i = \lambda_i/(N \cdot \lambda)$
  - *lloss*: try to achieve the same (local) loss rate at each site: $n_i = n'_i$ with $L = ErlangB(n'_i, \lambda_i/\mu)$
- Step 2: Inter-site rates depend on scheduling algorithm: always try local site (=job source) first, if busy
  - *SP*: shortest path (ie. closest free server)
  - *rand*: randomly pick a free site
  - *mostfree*: choose site with most free servers

Results

- Processing:
  - *mostfree* achieves highest local processing
  - intelligent server placement (prop, lloss) achieves higher local processing
- Link dimensions:
  - intelligent scheduling (*mostfree*) comes at a link bandwidth price





UNIVERSITEIT GENT

FACULTY OF ENGINEERING

Introduction
Network Architecture
Routing
Multiple Domains
Dimensioning
Conclusions

UNIVERSITEIT GENT

Conclusions

- **Optical Grid architecture**
  - OBS and/or hybrids seem very promising candidates
- **Routing**
  - Anycast routing requires deployment of new algorithms
- **Multiple Domains**
  - Proxy architecture to ensure scalability
- **Dimensioning**
  - Needs to be aligned with scheduling algorithms

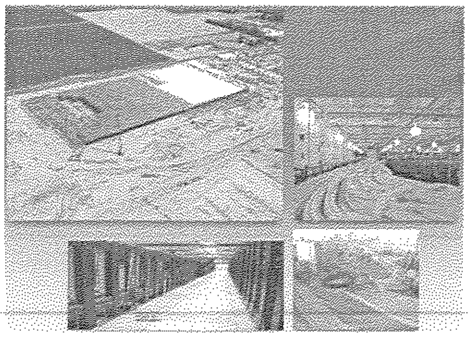
P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT p. 33 IBBT

Challenges

- **Integrated OCS/OBS/hybrid control plane**
  - Interworking, migration, node architecture, ..
- **Advanced dimensioning and network planning algorithms**
- **Resilience**
  - Job migration, protection/restoration approaches ..
- **Standardisation**
  - E.g. GoOBS architecture, burst format, routing protocols, inter-domain routing

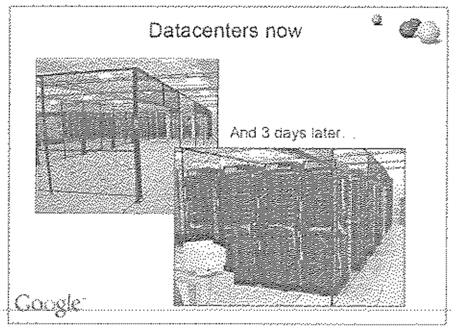
P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT p. 34 IBBT

Next ? Microsoft Server Plant



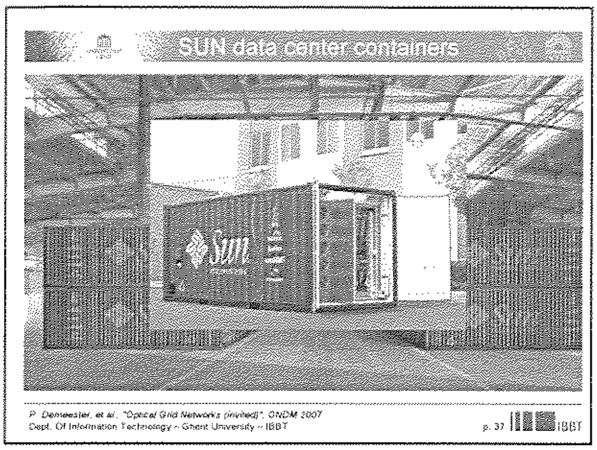
P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT p. 35 IBBT

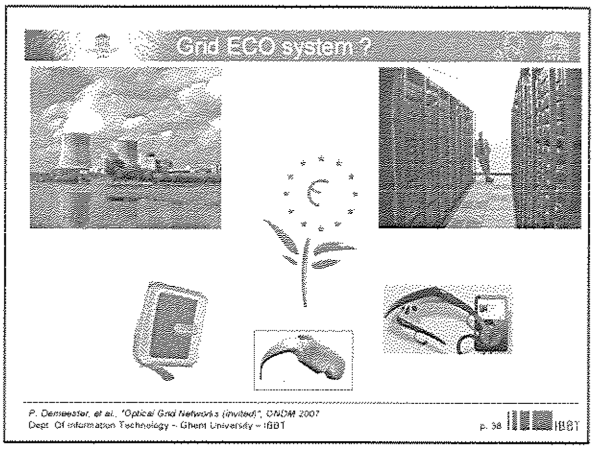
Google

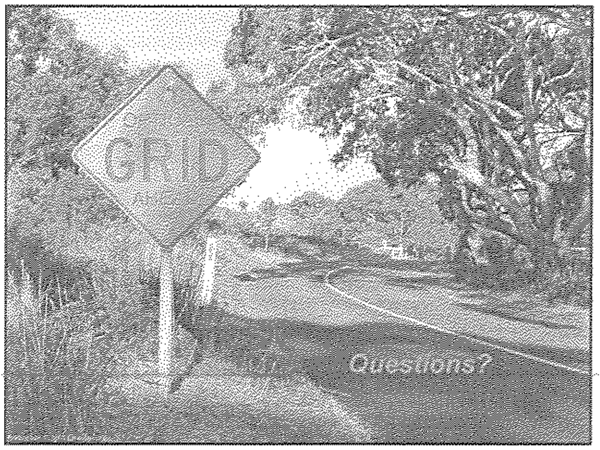


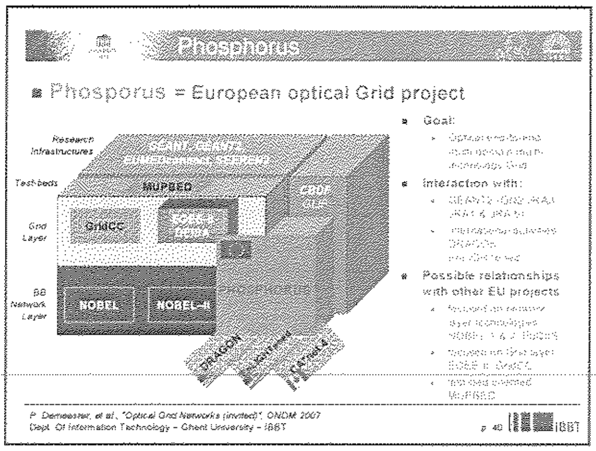
P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT p. 36 IBBT

UNIVERSITEIT GENT


SUN data center containers
Sun
P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT
p. 37
IBBT


Grid ECO system ?
P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT
p. 38
IBBT


GRID
Questions?


Phosphorus
Phosporus = European optical Grid project
Research Infrastructures
Test-beds
MUPBED
Grid Layer
GridCC
BB Network Layer
NOBEL
NOBEL-II
DRAGON
Goal:
Interaction with:
Possible relationships with other EU projects
P. Demeester, et al., "Optical Grid Networks (invited)", ONDM 2007
Dept. Of Information Technology – Ghent University – IBBT
p. 40
IBBT

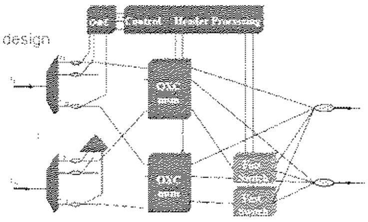
Phosphorus WP5

- Partners: AIT, CTI, UniBonn, i2Cat, IBBT/Ugent
- Supporting studies:
  - Job Demand modelling
  - Grid scheduling & routing
  - including QoS, multi-constraint routing (incl. physical impairments)
  - Advance Reservations
  - Control plane design
  - Simulation environment
  -

K. Christodoulopoulos et al., "Job Demand Models for Optical Grid Research", ONDM 2007

Hybrid switch architecture

- Joint work (ePhoton/ONe+) between UEssex and IBBT
- Work in progress:
  - Dimensioning
  - Scheduling
  - Node & network design

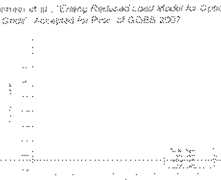


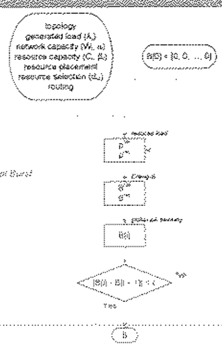
D. Simeonidou et al., "Design considerations for photonic routers supporting application-driven bandwidth reservations at sub-wavelength granularity", Proc. of WOBS'06

OBS Grid analysis

- Work in progress:
  - Performance modelling
    - Routing
    - Scheduling
  - Dimensioning
    - Resource placement
    - Capacity planning

M. De Leenheer et al., "Erlang Reduced Load Model for Optical Burst Switched Grids", Accepted for Proc. of GOBS 2007





UNIVERSITEIT GENT

Home

The 11th International Conference on Optical Networking Design and Modeling - ONDM 2007 - will be held in Athens, Greece on May 29-31, 2007.

ONDM conference is a major event in the rapidly growing area of optical networking addressing recent advances in the design, modeling and implementation of optical networks, including novel switching schemes and paradigms, network optimization and design, new concepts for link and control layer protocols, advanced network subsystems and node architectures, and network inter-working schemes.

A significant number of excellent invited speakers from leading academic and industrial institutions around the world will present their research work and view.

The ONDM 2007 conference program is now available.

Plenary Speakers:

Allan Willner (Univ. of Southern California)

Haruhisa Ichikawa (NTT - director)

Invited Speakers:

Keren Bergman (Columbia University)

Piero Castoldi (SSSUP)

Piet Demeester (GENT)

Andrea Fumagalli (UTD - Texas)

Maurice Gagnaire (GET)

Ken-ichi Kitayama (Osaka University)

Chunming Qiao (University at Buffalo)

Jian Wu (BUPT)

Hans Martin Foisel (T-Systems)

Dimitri Papadimitriou (Lucent-Alcatel)

A significant number of paper contributions from various research institutions around the world will also be presented. The full conference program will be posted once the final camera-ready versions of the accepted papers are received on March 23rd 2007.

The registration procedure will start on March 9th 2007.
Late registration will be available after April 30th 2007.

Please visit the relative sections on the left for more updates and information on registration, accommodation and the venue.

You can download the short version of call for participation here.

Sponsors

ONDM 2007

COST 291 2007 Workshop

Sponsors

ONDM 2007 is sponsored by the International Federation of Information Processing IFIP and supported by:

the EU ePhoton/ONe project (http://www.e-photon-one.org/)

and COST 291 action (http://www.ait.edu.gr/cost291/).

Organizer

ATHENS INFORMATION TECHNOLOGY

Athens Information Technology (AIT)
P.O. Box 68
Markopoulo Ave.
GR - 19002 Peania
Athens (Greece)